RSAC 2025 Conference

Many Voices.
One Community.

SESSION ID: HTA-T01

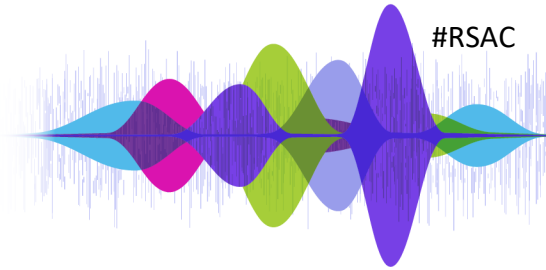
Cloud, Identity & SaaS Forensic Investigation—

Not What You Think!

Ofer Maor
CTO, Mitiga

linkedin.com/in/ofermaor/

MITIGA



Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the presenters individually and, unless expressly stated to the contrary, are not the opinion or position of RSA Conference LLC or any other co-sponsors. RSA Conference LLC does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented.

Attendees should note that sessions may be audio- or video-recorded and may be published in various media, including print, audio and video formats without further notice. The presentation template and any media capture are subject to copyright protection.

© 2025 RSA Conference LLC or its affiliates. The RSAC and RSAC CONFERENCE logos and other trademarks are proprietary. All rights reserved.

CloudSec & AppSec
Daytime

Incident Response
Nights & Weekends

Over 30 Years in Security (25 in Prevention)!

Builder, Breaker, Defender

Pioneer of IAST

OWASP Board & Contributor

Hacker at Heart

Cloud & SaaS Threat Detection, Investigation & Response

Pioneer of CIRA

CTO & Co-Founder MITIGA

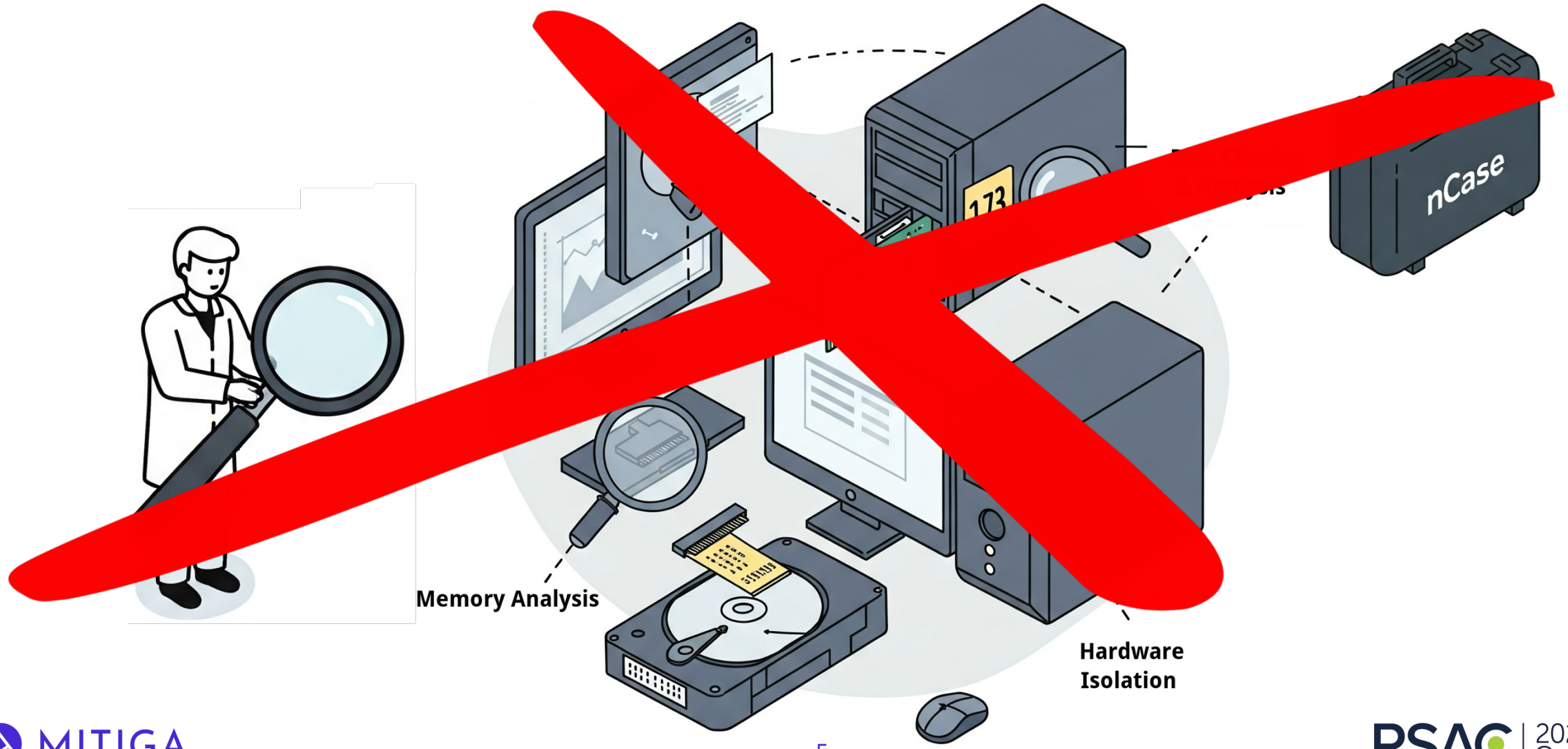
RSAC 2025 Conference

Introduction

Cloud (& SaaS) Forensics Overview

Many Voices.
One Community.

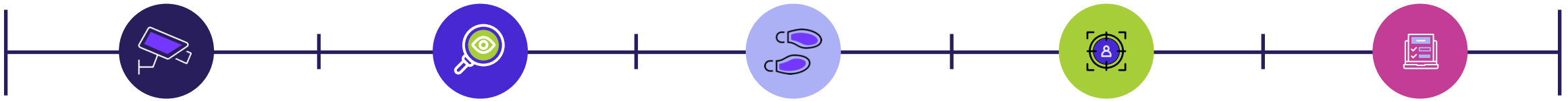
What Is Forensics Investigation?



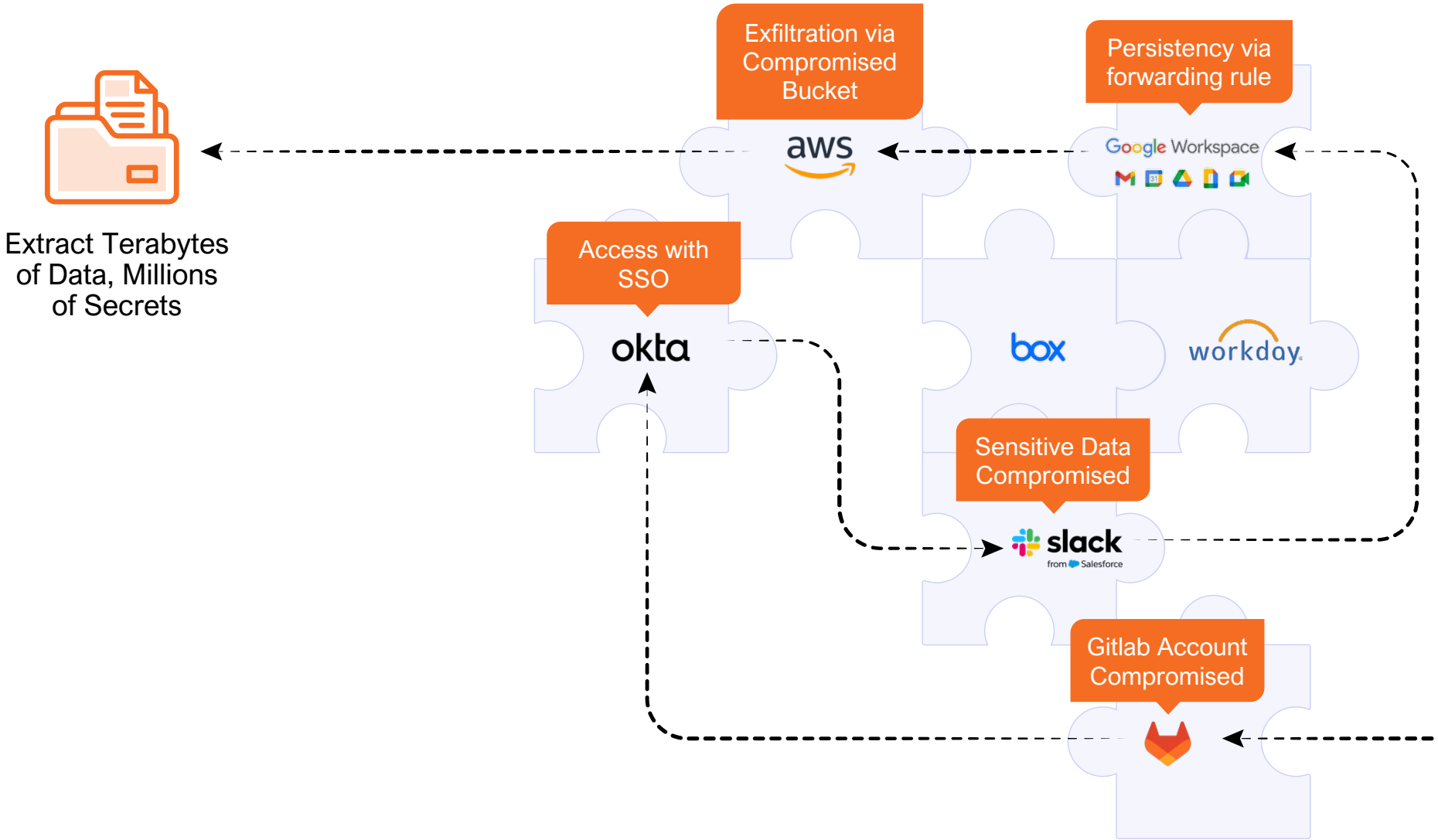
What Is Forensics Investigation?

A **forensic investigation** is the process of gathering, analyzing, and preserving evidence to uncover the facts behind a crime, dispute, or suspicious incident.

Words of Wisdom by ChatGPT



A Typical Cloud Attack



Your Cloud

Cloud Infrastructure

aws

Made up of many services:

Workloads / Storage / Network/ Database / Security

Cloud Corporate IT (Identity, Office, etc.)

Azure Active Directory

Office 365

Google Workspace

okta

SaaS Applications

1Password

Jira Software

slack

workday.

mongoDB.

snowflake

salesforce

GitLab

ATLASSIAN

LastPass

box

Traditional DFIR (Host Based)

Cloud Infrastructure

aws

Made up of many services:

Workloads / Storage / Network/ Database / Security

Cloud Corporate IT (Identity, Office, etc.)

Azure Active Directory

Office 365

Google Workspace

okta

SaaS Applications

1Password

Jira Software

slack

workday.

mongoDB.

snowflake

salesforce

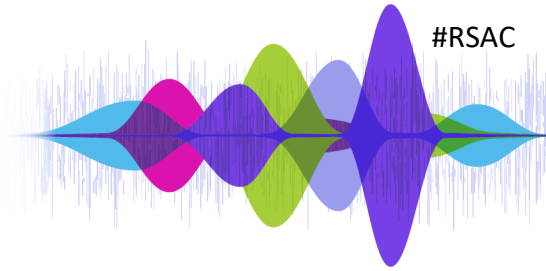
GitLab

ATLASSIAN

LastPass

box

Cloud Telemetray Challenges



Security Logging *(vs. Activity)*	Bad Defaults	Content (Missing Events, Fields)
Variance (Structure, Fields)	Configuration	Timing
Retention	Large Volume	Cost (SIEM, egress, generate)

No Default Access Logs for CSPs

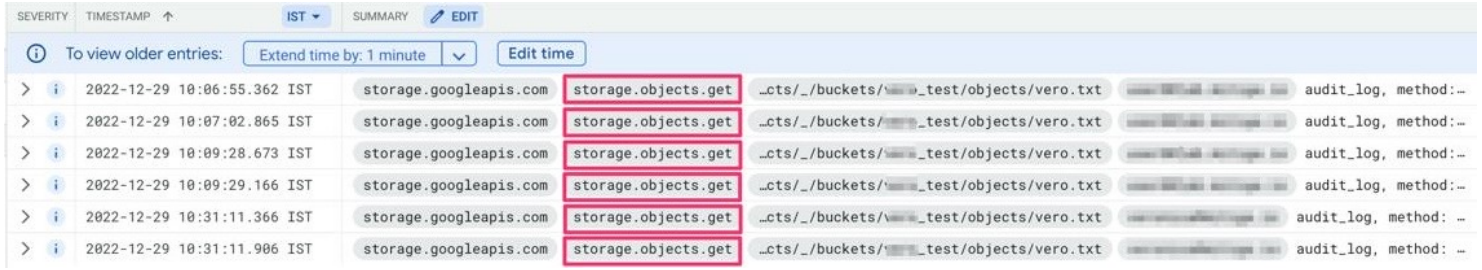
#RandomExamples

By default, Amazon S3 doesn't collect server access logs. When you enable logging, Amazon S3 delivers access logs for a source bucket to a destination bucket (also known as a *target bucket*) that you choose. The destination bucket must be in the same AWS Region and AWS account as the source bucket.

- Data Access audit logs—except for BigQuery—are disabled by default. If you want Data Access audit logs to be written for Google Cloud services other than BigQuery, you must explicitly enable them.
- Data Access audit logs are stored in the `_Default` bucket unless you've routed them elsewhere. For more information, see Storing and routing audit logs.
- Data Access audit logs help Google Support troubleshoot issues with your account. Therefore, we recommend enabling Data Access audit logs when possible.

GCP Limited Data Access Information

#RandomExamples



SEVERITY | TIMESTAMP | IST | SUMMARY | EDIT

To view older entries: Extend time by: 1 minute | Edit time

2022-12-29 10:06:55.362 IST	storage.googleapis.com	storage.objects.get	…cts/_/buckets/…_test/objects/vero.txt	[redacted]	audit_log, method:…
2022-12-29 10:07:02.865 IST	storage.googleapis.com	storage.objects.get	…cts/_/buckets/…_test/objects/vero.txt	[redacted]	audit_log, method:…
2022-12-29 10:09:28.673 IST	storage.googleapis.com	storage.objects.get	…cts/_/buckets/…_test/objects/vero.txt	[redacted]	audit_log, method:…
2022-12-29 10:09:29.166 IST	storage.googleapis.com	storage.objects.get	…cts/_/buckets/…_test/objects/vero.txt	[redacted]	audit_log, method:…
2022-12-29 10:31:11.366 IST	storage.googleapis.com	storage.objects.get	…cts/_/buckets/…_test/objects/vero.txt	[redacted]	audit_log, method: …
2022-12-29 10:31:11.906 IST	storage.googleapis.com	storage.objects.get	…cts/_/buckets/…_test/objects/vero.txt	[redacted]	audit_log, method: …

Event type	GCP	AWS
Read the Metadata of an object	storage.objects.get	GetObject + sub field:response-content-disposition
Read the content of a file	storage.objects.get	GetObject + sub field:response-content-disposition:inline
Download a file locally	storage.objects.get	GetObject + sub field:response-content-disposition:attachment
Copy file to internal bucket	storage.objects.get + storage.objects.create	CopyObject + sub fields: account id and bucket ARN of the internal destination bucket.
Export file to external bucket	storage.objects.get	CopyObject + sub fields: account id and bucket ARN of the destination bucket.

MITIGA

AWS Logs Assumed Role Identity

#RandomExamples

```
"userIdentity": {
    "type": "AssumedRole",
    "arn": "arn:aws:sts::012345678910:assumed-role/my-airflow-role-1c34dg",
    "accountId": "012345678910",
    "accessKeyId": "ASIAQV9HXLTLJIT5R3XCH",
    "sessionContext": {
        "sessionIssuer": {
            "type": "Role",
            "arn": "arn:aws:iam::012345678910:role/my-airflow-role-1c34dg",
            "accountId": "012345678910",
            "userName": "my-airflow-role-1c34dg"
        },
        "webIdFederationData": {
            "federatedProvider": "arn:aws:iam::012345678910:oidc-provider/oidc.eks.us-east-1
        }
    }
```

Salesforce Log User Masking

#RandomExamples

USER_ID

Type

Id

Description

The 15-character case insensitive ID of the user who's using Salesforce services through the UI or the API.

For example: `00590000000I1SNIA0`.

Snowflake Logging Is a Database

#RandomExamples

Snowflake logs are stored in default database called **SNOWFLAKE** *inside the* **ACCOUNT_USAGE** *schema.*

*This provides a comprehensive (**yet uncommon**) audit trail.*

ACCOUNT_USAGE

- QUERY_HISTORY
- LOGIN_HISTORY
- SESSIONS
- ACCESS_HISTORY

Cloud Telemetray Challenges

Security Logging *(vs. Activity)* ***Need to Activate***	Bad Defaults ***Need to Activate***	Content (Missing Events, Fields) ***Enrichment, Telemetry***
Variance (Structure, Fields) ***Normalization, Organization***	Configuration ***Enrichment & Context***	Timing ***Logging Tiers, Alternates***
Retention ***Collection & Storage***	Large Volume ***Data Warehousing***	Cost (SIEM, egress, generate) ***Data Warehousing***

RSAC 2025 Conference

Obtaining Forensics Data

How to Collect & Store Forensics Telemetry

Many Voices.
One Community.

Logs??? I Have a SIEM!

Well,

Good Luck!

- Not designed for scale and complexity of cloud telemetry
- Query language/capabilities not optimized for cloud complexity
- SIEM cost
- Cloud egress cost

The Data Lake (Data Warehouse) Approach

- Designed for Big Data applications
- Supports SQL querying as well as advanced coding (e.g. PySpark)
- Normalization, enrichment & aggregation of data
- Can be distributed, kept *at source* (no egress cost)

Ultimately - provides a more suitable framework for cloud logging, analytics & forensics investigation

Creating Your Own Data Lake

DuckDB

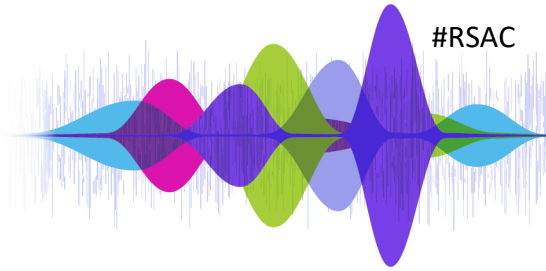
MITIGA

```
import time
import duckdb

cores = 1
con = duckdb.connect() # Connect to DuckDB (in-memory database)
print("connected to duckdb")

con.execute(f"SET threads={cores}")
con.execute("SET memory_limit='1GB'")
con.execute("INSTALL DELTA")
con.execute("LOAD DELTA")
start_time = time.perf_counter()

input_path_gz = "./data/00_00_00_00_59_59_S0.json.gz"
input_path = input_path_gz
output_path = "/tmp/output_duck_parquet"

# SQL to extract and optimize common audit log fields
sql = """
    COPY ( SELECT * FROM read_json_auto(
            '{input_path}', format='newline_delimited', auto_detect=false,
            COLUMNS = {{insertId: 'VARCHAR', labels: 'VARCHAR', logName: 'VARCHAR',
                operation: 'STRUCT(id VARCHAR, producer VARCHAR, first BOOLEAN, last BOOLEAN)',
                protoPayload: 'STRUCT(callerIp VARCHAR, principalEmail VARCHAR, serviceAccountKeyName VARCHAR, project_id VARCHAR,
                    serviceName VARCHAR, methodName VARCHAR, resourceName VARCHAR, status VARCHAR, metadata VARCHAR,requestMetadata VARCHAR,
                    request VARCHAR, authenticationInfo STRUCT(principalEmail VARCHAR, principalSubject VARCHAR), response VARCHAR
                )',
                receiveTimestamp: 'VARCHAR', resource: 'VARCHAR', severity: 'VARCHAR', timestamp: 'VARCHAR', httpRequest: 'VARCHAR'
            }}, maximum_depth=2
    ))
    TO '{output_path}'
    (   FORMAT PARQUET,
        COMPRESSION 'ZSTD',
        ROW_GROUP_SIZE 122880
    )
""".format(input_path=input_path, output_path=output_path)

con.execute(sql)
end_time = time.perf_counter()
total_rows = 1976486
actual_count = duckdb.sql(f"SELECT COUNT(*) FROM read_parquet('{output_path}')").fetchone()[0]
assert actual_count==total_rows
print(f"Total time: {end_time-start_time}. EPS {total_rows/(end_time-start_time)}, per core: {total_rows/(end_time-start_time)/cores}")
```

```
bq load \
  --source_format=NEWLINE_DELIMITED_JSON \
  --autodetect \
  your_dataset.your_table \
  gs://your_bucket/your_file.json
```

Or just use BigQuery

Inferred vs Defined SCHEMA



Inferred

- Auto schema on existing dataset
- Easy, automated
- Doesn’t scale over time (data may change)
- Harder to investigate on
- Slower (more expensive)

Defined

- Based on a predefined schema
- Requires providing a schema
- Better support for various types of data (and future changes)
- Optimized for investigations
- More efficient (speed, cost)

GuardDuty Example


```
"Id": "aecaea323076448bde53a8c0bc4cb496",
"Partition": "aws",
"Region": "us-east-1",
"Resource": {
    "AccessKeyDetails": {
        "AccessKeyId": "ASIA5FW4NLJMYHDY3XLL",
        "PrincipalId": "902449150213",
        "UserName": "alice-acts-dev",
        "UserType": "Root"
    },
    "ResourceType": "AccessKey"
},
"SchemaVersion": "2.0",
"Service": {
    "Action": {
        "ActionType": "AWS_API_CALL",
        "AwsApiCallAction": {
            "AffectedResources": {},
            "Api": "GetFeatureOptInStatus",
            "CallerType": "Remote IP",
            "ErrorCode": "AccessDenied",
            "RemoteIpDetails": {…
            },
            "ServiceName": "notifications.amazonaws.com"
        }
    },
```

```
        "KubernetesUserDetails": {
            "Groups": [
                "system:serviceaccounts",
                "system:serviceaccounts:samplebob",
                "system:authenticated"
            ],
            "Uid": "ac29dc22-c0dd-2683-abcd-e485712
            "Username": "system:serviceaccount:sampl
        },
        "KubernetesWorkloadDetails": {
            "HostNetwork": "false",
            "Name": "karpenter-controller-helm-5b9e
            "Namespace": "kube-system",
            "Type": "pods",
            "Uid": ""
        }
    },
    "ResourceType": "EKSCluster"
},
"SchemaVersion": "2.0",
"Service": {
    "Action": {
        "ActionType": "KUBERNETES_API_CALL",
        "KubernetesApiCallAction": {
            "RemoteIpDetails": {…
            },
            "RequestUri": "/api/v1/namespaces/kube-s
            "StatusCode": "101",
            "UserAgent": "Go-http-client/1.1",
            "Verb": "create"
```

```
"Tags": [
    {
        "Key": "usage-id",
        "Value": "AA00002473"
    },
    {
        "Key": "sd-period",
        "Value": "NA"
    },
    {
        "Key": "exp-date",
        "Value": "99-00-9999"
    },
    {
        "Key": "alpha.alice.io/cluster-name",
        "Value": "mcp-msd-use1-dev-ns-01"
    },
    {
        "Key": "alpha.alice.io/stack-uuid",
        "Value": "mt14q0;-cluster-4b3e3601-3f7a-4b1b-
    },
    {
        "Key": "env-type",
        "Value": "dev"
    },
    {
        "Key": "alpha.alice.io/alice-version",
        "Value": "v2.5.0-1"
    },
    {
        "Key": "cost-center",
        "Value": "522062"
    }
```

GuardDuty – Log Complexity

1. **Comprehensive Coverage**
   Multiple logged resources: `InstanceDetails`, `KuberenetesDetails`, `ECSClusterDetails`, `S3BucketDetails`, etc.

2. **Rich Action Tracking**
   Supports multiple event: `AwsApiCallAction`, `DNSRequestAction`, `KubernetesApiCallAction`, `PortProbeAction`, etc.

3. **Detailed Geographic Information**
   Extensive geographic tracking: `CityName`, `CountryName:CountryCode`, `GeoLocation-Lat:Lon`, etc.

4. **Deep Network Analysis**
   Tracks both IPv4 & IPv6, Network interfaces, VPC details, ISP/ASN, port-level details, etc.

5. **Threat Intelligence Integration**
   Supports adding TI-related fields

6. **And More…**

Sample Defined GuardDuty Schema

```
guardDutySchema = """{"fields":[{"metadata":{},"name":"AccountId","nullable":true,"type":"string"},{"metadata":{},
"name":"Arn","nullable":true,"type":"string"},{"metadata":{},"name":"Confidence","nullable":true,"type":"string"},
{"metadata":{},"name":"CreatedAt","nullable":true,"type":"string"},{"metadata":{},"name":"Description","nullable":true,
"type":"string"},{"metadata":{},"name":"Id","nullable":true,"type":"string"},{"metadata":{},"name":"Partition",
"nullable":true,"type":"string"},{"metadata":{},"name":"Region","nullable":true,"type":"string"},{"metadata":{},
"name":"Resource","nullable":true,"type":{"fields":[{"metadata":{},"name":"AccessKeyDetails","nullable":true,"type":
{"fields":[{"metadata":{},"name":"AccessKeyId","nullable":true,"type":"string"},{"metadata":{},"name":"PrincipalId",
"nullable":true,"type":"string"},{"metadata":{},"name":"UserName","nullable":true,"type":"string"},{"metadata":{},
"name":"UserType","nullable":true,"type":"string"}],"type":"struct"}},{"metadata":{},"name":"ContainerDetails",
"nullable":true,"type":{"fields":[{"metadata":{},"name":"ContainerRuntime","nullable":true,"type":"string"},{"metadata":{},
"name":"Id","nullable":true,"type":"string"},{"metadata":{},"name":"Image","nullable":true,"type":"string"},{"metadata":{},
"name":"ImagePrefix","nullable":true,"type":"string"},{"metadata":{},"name":"Name","nullable":true,"type":"string"},
{"metadata":{},"name":"SecurityContext","nullable":true,"type":"string"},{"metadata":{},"name":"VolumeMounts",
"nullable":true,"type":"string"}],"type":"struct"}},{"metadata":{},"name":"ResourceType","nullable":true,"type":"string"}],
"type":"struct"}},{"metadata":{},"name":"SchemaVersion","nullable":true,"type":"string"},{"metadata":{},"name":"Service",
"nullable":true,"type":{"fields":[{"metadata":{},"name":"Action","nullable":true,"type":{"fields":[{"metadata":{},
"name":"ActionType","nullable":true,"type":"string"},{"metadata":{},"name":"ServiceName","nullable":true,
"type":"string"}]}},{"metadata":{},"name":"DetectorId","nullable":true,"type":"string"},{"metadata":{},
"name":"EventFirstSeen","nullable":true,"type":"string"},{"metadata":{},"name":"EventLastSeen","nullable":true,
"type":"string"}]}},{"metadata":{},"name":"Severity","nullable":true,"type":"string"},{"metadata":{},"name":"Title",
"nullable":true,"type":"string"},{"metadata":{},"name":"Type","nullable":true,"type":"string"},{"metadata":{},
"name":"UpdatedAt","nullable":true,"type":"string"}],"type":"struct"}"""
```

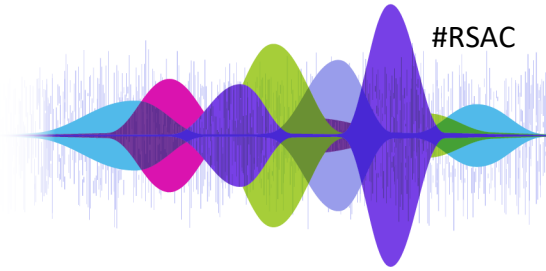
Open Cybersecurity Schema Framework

- Standardized schema for Events
- Driven by AWS, now part of Linux Foundation
- Limited functionality
- Limited adoption

1.4k followers

Collecting the Right Telemetry

Partial Lists



aws

- Metadata (Configuration)
- CloudTrail
- GuardDuty
- S3 Object Level
- DNS (Route53) Logs
- VPC Flow Logs
- ELB Logs
- EKS Control Plane Logs

Google Cloud Platform

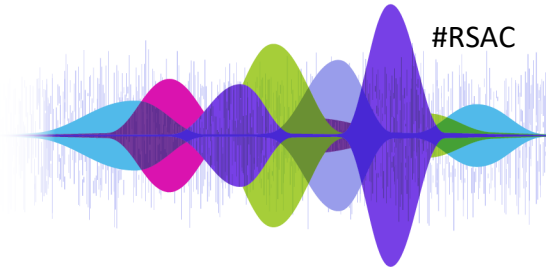
- Metadata (Configuration)
- Activity Audit Logs
- Data Access Audit Logs
- VPC Flow Logs
- DNS Logs
- Load Balancer Logs

Azure

- Metadata of Tenant
- Metadata per Subscription
- Activity Logs
- Blob Storage Logs
- AKS Logs
- Application Gateway Logs
- NSG Flow Logs

Collecting the Right Telemetry

Partial Lists



- Azure AD Unified Audit Logs
- Azure AD Audit
- General Audit Log
- SharePoint Audit Log
- Data Loss Prevention Audit Log
- Exchange Audit Log
- Email Flow Logs
- Interactive User Sign-ins
- Non-interactive User Sign-ins
- Managed Identity Sign-ins
- Service Principal Sign-ins
- Directory Audit Log
- Microsoft Graph Activity Log

- Admin Activity Logs
- Login Activity Logs
- User Accounts Logs
- SAML Logs
- Groups Logs
- Drive Activity Log
- GCP Log
- Groups Enterprise Log
- Token Log
- Chrome Log
- Mobile Log
- Context-Aware Access Log
- Calendar Log
- Chat Log
- Meet Log
- Rules Log

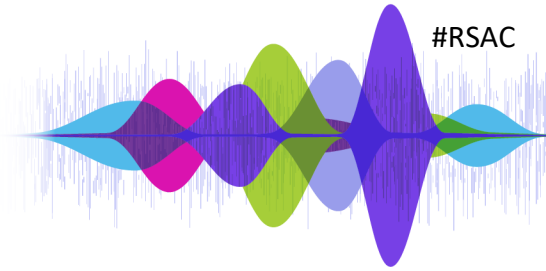
RSAC 2025 Conference

Investigation

Cloud & SaaS Forensics Investigation Examples

Many Voices.
One Community.

Querying the Data?



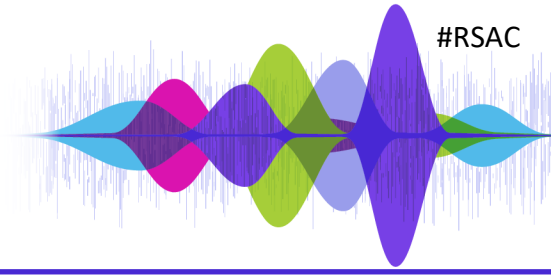
SQL

- Declarative, straightforward for standard querying
- Optimized for structured data with good indexes
- Simple and intuitive. Ideal for ad-hoc analysis
- Mainly used for data extraction

PySpark

- More expressive, allows for complex manipulations
- Distributed computing with improved performance
- More complex. Based on Python and Spark APIs
- More suitable for data modifications (as well as ML)

Salesforce Log User Masking



USER_ID

Type

Id

Description

The 15-character case insensitive ID of the user who's using Salesforce services through the UI or the API.

For example: `00590000000I1SNIA0`.

Salesforce Query API

Developers

Syntax

URI

`/services/data/vXX.X/query?q=query`

Formats

JSON, XML

HTTP Method

GET

Authentication

`Authorization: Bearer token`

Requires Auth.
Requires Code

```
{{_endpoint}}/services/data/v{{version}}/query/?q=SELECT
UserName,Id,Alias,Email,UserType,CreatedDate,IsActive,
LastLoginDate,LastPasswordChangeDate,Title,UserRoleId FROM User
```


```
{
    "attributes": {
        "type": "User",
        "url": "/services/data/v57.0/sobjects/User/0058e000001NfTcJJs"
    },
    "Username": "john@domain.com",
    "Id": "0058e000001NfTcJJs",
    "Alias": "john",
    "Email": "john@domain.com",
    "UserType": "User",
    "CreatedDate": "2024-05-09T11:56:14.000+0000",
    "IsActive": true,
    "LastLoginDate": null,
    "LastPasswordChangeDate": null,
    "Title": null,
    "UserRoleId": null
}
```

Workaround – Enrich From Logs

```
def add_user_name_to_event_log_files(EventLogFiles_df):
    logins_df = EventLogFiles_df\
        .filter(col('EVENT_TYPE') == 'Login')\
        .filter(col('user_name').isNotNull())\
        .filter(col('USER_ID').isNotNull())\
        .groupBy('ORGANIZATION_ID', 'date', 'USER_ID').agg(_min('TIMESTAMP_DERIVED').alias
        ('mtg_login_time'), collect_set(lower('user_name')).alias('mtg_user_name'))\
        .withColumn('mtg_user_name', when(size(col('mtg_user_name')) > 1, col('mtg_user_name').
        cast(StringType())).otherwise(col('mtg_user_name').getItem(0)))

    enriched_df = EventLogFiles_df.join(logins_df, on=['ORGANIZATION_ID', 'date', 'USER_ID'],
    how='left')

    return enriched_df
```

Using It in Practice

```
def agg_huge_amount_of_query_results(raw_salesforce_eventlogfiles_df: DataFrame, threshold: int = 1024*1024*10,
time_window:str="24 hours", *args, **kwargs) -> DataFrame:
    """ Description: Identifies big amount of query results (in bytes).
    The function returns a df containing all the relevant events. """

    enriched_df = add_user_name_to_event_log_files(raw_salesforce_eventlogfiles_df)
    filtered_df = enriched_df\
        .filter(col('EVENT_TYPE') == EVENT_NAME)\
        .withColumn(COL_URI, lower(get_json_object('json', '$.URI')))\
        .filter(col(COL_URI).contains('/query/') | col(COL_URI).contains('/queryall/'))\
        .withColumn(COL_RESPONSE_SIZE, get_json_object('json', '$.RESPONSE_SIZE'))\
        .filter(col(COL_RESPONSE_SIZE).cast(IntegerType()) > threshold)\
        .filter(get_json_object('json', '$.METHOD') == 'GET')\
        .filter(get_json_object('json', '$.STATUS_CODE') == '200')\
        .filter(get_json_object('json', '$.REQUEST_STATUS') == 'S')\
        .withColumn(COL_ENTITY_NAME, get_json_object('json', '$.ENTITY_NAME'))

    # Aggregating on time window
    agg_df = filtered_df.withColumn("temp_time_window", (window('TIMESTAMP_DERIVED', time_window).start))
```

Google Workspace – Logs4Pros

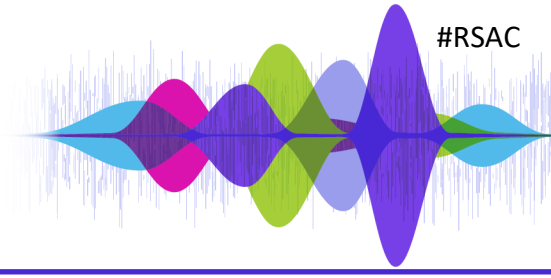
```
"actor": {
    "callerType": "KEY",
    "email": null,
    "key": "SYSTEM",
    "profileId": null
},
"etag": "\"7gJ2FxxdqGNyNxXZrlQg-FP23AygT_AilQ6bYcusikU/9bzgsFGI9nGOlALDpadlFaRF53Y\"",
"events": [
    {
        "name": "USER_LICENSE_ASSIGNMENT",
        "parameters": [
            {
                "name": "USER_EMAIL",
                "value": "john.doe@demogcp.com"
            },
            {
                "name": "PRODUCT_NAME",
                "value": "Google Workspace Archived User"
            },
            {
                "name": "NEW_VALUE",
                "value": "Google Workspace Enterprise Plus - Archived User"
            }
        ],
        "type": "LICENSES_SETTINGS"
    },
    {
        "name": "USER_LICENSE_REVOKE",
        "parameters": […
        ],
        "type": "LICENSES_SETTINGS"
    }
```


Multiple events clumped together *(same actor)*

Names of fields are actually ***values***

Parameters vary, difficult to search on

Parameters, Parameters, Parameters



"user": "john"

↓

"name": "user"
"value": "john"

```
"name": "MOVE_USER_TO_ORG_UNIT",
"parameters": [
    {
        "name": "USER_EMAIL",
        "value": "john.doe@demogcp.com"
    },
    {
        "name": "ORG_UNIT_NAME",
        "value": "/"
    },
    {
        "name": "NEW_VALUE",
        "value": "/OU-Archived Contractors"
    }
```

```
"name": "USER_LICENSE_ASSIGNMENT",
"parameters": [
    {
        "name": "USER_EMAIL",
        "value": "john.doe@demogcp.com"
    },
    {
        "name": "PRODUCT_NAME",
        "value": "Google Workspace Archived User"
    },
    {
        "name": "NEW_VALUE",
        "value": "Google Workspace Enterprise Plus - Archived User"
    }
```

```
"name": "SECURITY_CENTER_RULE_THRESHOLD_TRIGGER",
"parameters": [
    {
        "name": "SECURITY_CENTER_RULE_NAME",
        "value": "User License's revoked and reassigned"
    },
    {
        "name": "SECURITY_CENTER_RULE_TRIGGER_WINDOW",
        "value": "PT24H"
    },
    {
        "name": "SECURITY_CENTER_RULE_CONDITION",
        "value": "(EVENT EQUALS [ADMIN_EVENTS_USER_LICENSE_REVOKE])
    },
    {
        "name": "SECURITY_CENTER_RULE_THRESHOLD",
        "value": "COUNT > 2"
    },
    {
        "name": "SECURITY_CENTER_RULE_TIME_FRAME",
        "value": "2025-03-24T19:44:00Z - 2025-03-25T19:45:00Z"
    },
    {
        "name": "SECURITY_CENTER_RULE_ACTION",
        "value": null
    }
```

Usable GWS Logs – The Magic

```
if('parameters' in df_events.schema.simpleString()):
    df_events = df_events.withColumn('parameters_str', col('col.parameters')).drop('col')

    param_list = df_events.schema['parameters_str'].dataType.elementType.fieldNames()

    cast_expr_inner = ', '.join([f'{param}: STRING' for param in param_list])
    cast_expr = f"CAST(parameters_str AS ARRAY<STRUCT<{cast_expr_inner}>>)"
    df_events = df_events.withColumn("parameters_str",  expr(cast_expr))

    param_values_expr_inner = ', '.join([f'x.{param}' for param in param_list if param != 'name'])
    param_values_expr = f"transform(parameters_str, x -> filter(array({param_values_expr_inner}), y -> y IS NOT NULL)[0])"
    df_events = df_events.withColumn("param_values", expr(param_values_expr))

    df_events = df_events.withColumn("parameters", map_from_arrays("parameters_str.name", "param_values"))
    df_events = df_events.withColumn('parameters', when(col('parameters').isNull(), create_map(lit("isNull"), lit(True).cast
    (StringType()))).otherwise(col('parameters')))

    df = df_events.drop('parameters_str', 'param_values')
else:
    df = df_events.drop('col')
```

Google Workspace – The Usable Version

actor	name	type	parameters
{"callerType":"USER","email":"johnd	ALERT_CENTER_VIEW	ALERT_CENTER	{"ALERT_ID":"8209188c-5786-4f1a-8284-3c6fe4b63180,c690ed69-101a-4cba-bf1f-ed5430c1fdb9,cff71906-7561-4a5a-b08b-(
{"callerType":"USER","email":"bob@	CREATE_DATA_TRANSFER_REQ	USER_SETTINGS	{"USER_EMAIL":"alice@gcpdemo.com","DESTINATION_USER_EMAIL":"bob@gcpdemo.com","APPLICATION_NAME":"Drive and
{"callerType":"USER","email":"alice(	ALERT_CENTER_VIEW	ALERT_CENTER	{"ALERT_ID":"8209188c-5786-4f1a-8284-3c6fe4b63180,a412c232-a09b-43e5-8013-3443aa48a651,cfa3d1af-7579-4e86-9992
{"callerType":"USER","email":"alice(	ALERT_CENTER_VIEW	ALERT_CENTER	{"ALERT_ID":"8209188c-5786-4f1a-8284-3c6fe4b63180,a412c232-a09b-43e5-8013-3443aa48a651,cfa3d1af-7579-4e86-9992
{"callerType":"USER","email":"alice(	ALERT_CENTER_VIEW	ALERT_CENTER	{"ALERT_ID":"8209188c-5786-4f1a-8284-3c6fe4b63180,a412c232-a09b-43e5-8013-3443aa48a651,cfa3d1af-7579-4e86-9992
{"callerType":"USER","email":"okta.g	SUSPEND_USER	USER_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com"}
{"callerType":"KEY","email":null,"ke	REMOVE_GROUP_MEMBER	GROUP_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com","GROUP_EMAIL":"nycore@gcpdemo.com"}
{"callerType":"KEY","email":null,"ke	REMOVE_GROUP_MEMBER	GROUP_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com","GROUP_EMAIL":"employees@gcpdemo.com"}
{"callerType":"KEY","email":null,"ke	REMOVE_GROUP_MEMBER	GROUP_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com","GROUP_EMAIL":"usg@gcpdemo.com"}
{"callerType":"KEY","email":null,"ke	MOVE_USER_TO_ORG_UNIT	USER_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com","ORG_UNIT_NAME":"/","NEW_VALUE":"/OU-Arvhived employees"}
{"callerType":"KEY","email":null,"ke	ARCHIVE_USER	USER_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com"}
{"callerType":"KEY","email":null,"ke	USER_LICENSE_ASSIGNMENT	LICENSES_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com","PRODUCT_NAME":"Google Workspace Archived User","NEW_VALUE":"Google W
{"callerType":"KEY","email":null,"ke	USER_LICENSE_REVOKE	LICENSES_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com","PRODUCT_NAME":"Google Workspace","OLD_VALUE":"Google Workspace Enter
{"callerType":"USER","email":"oktaa	SUSPEND_USER	USER_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com"}
{"callerType":"KEY","email":null,"ke	MOVE_USER_TO_ORG_UNIT	USER_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com","ORG_UNIT_NAME":"/","NEW_VALUE":"/OU-Arvhived employees"}
{"callerType":"KEY","email":null,"ke	ARCHIVE_USER	USER_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com"}
{"callerType":"KEY","email":null,"ke	USER_LICENSE_ASSIGNMENT	LICENSES_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com","PRODUCT_NAME":"Google Workspace Archived User","NEW_VALUE":"Google W
{"callerType":"KEY","email":null,"ke	USER_LICENSE_REVOKE	LICENSES_SETTINGS	{"USER_EMAIL":"johndemo@gcpdemo.com","PRODUCT_NAME":"Google Workspace","OLD_VALUE":"Google Workspace Enter
{"callerType":"KEY","email":null,"ke	USER_LICENSE_ASSIGNMENT	LICENSES_SETTINGS	{"USER_EMAIL":"elizabeth@gcpdemo.com","PRODUCT_NAME":"Google Workspace","NEW_VALUE":"Google Workspace Enterp
{"callerType":"USER","email":"oktaa	CREATE_USER	USER_SETTINGS	{"USER_EMAIL":"elizabeth@gcpdemo.com"}
{"callerType":"KEY","email":null,"ke	USER_LICENSE_ASSIGNMENT	LICENSES_SETTINGS	{"USER_EMAIL":"albert@gcpdemo.com","PRODUCT_NAME":"Google Workspace","NEW_VALUE":"Google Workspace Enterpris
{"callerType":"KEY","email":null,"ke	USER_LICENSE_ASSIGNMENT	LICENSES_SETTINGS	{"USER_EMAIL":"michaela@gcpdemo.com","PRODUCT_NAME":"Google Workspace","NEW_VALUE":"Google Workspace Enter
{"callerType":"KEY","email":null,"ke	USER_LICENSE_ASSIGNMENT	LICENSES_SETTINGS	{"USER_EMAIL":"test@gcpdemo.com","PRODUCT_NAME":"Google Workspace","NEW_VALUE":"Google Workspace Enterprise F
{"callerType":"USER","email":"oktaa	CREATE_USER	USER_SETTINGS	{"USER_EMAIL":"albert@gcpdemo.com"}
{"callerType":"USER","email":"oktaa	CREATE_USER	USER_SETTINGS	{"USER_EMAIL":"michaela@gcpdemo.com"}

The Azure Multiplicity

Single Azure operation manifests as multiple log events, separated by *stages* of operation:

`Started, Accepted, Succeeded, Etc.`

category	correlationId	operationName	properties	result	time
Administrative	b14b721d-9ce3-44f9	MICROSOFT.COMPUTE/VIRTUALMACHINES/WRITE	{"eventCategory":"Administrat	Started.	2025-03-27T09:18:31.22
Administrative	b14b721d-9ce3-44f9	MICROSOFT.COMPUTE/VIRTUALMACHINES/WRITE	{"statusCode":"Created","serv	Accepted.Created	2025-03-27T09:18:33.31
Administrative	b14b721d-9ce3-44f9	MICROSOFT.COMPUTE/VIRTUALMACHINES/WRITE	{"eventCategory":"Administrat	Succeeded.	2025-03-27T09:20:19.88

Information relevant for investigation (or detection) is split across the events

result also provides insight to *operationName*: WRITE can be WRITE **or** UPDATE

Azure Multiplicity Example

VM Creation

category	correlationId	operationName	properties	result	resultType	time
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	{"eventCategory":"Administrative","entity":"/subscript	Started.	Start	2025-03-27T09:18:31.22
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	{"statusCode":"Created","serviceRequestId":null,"res	Accepted.Created	Accept	2025-03-27T09:18:33.31
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	{"eventCategory":"Administrative","entity":"/subscript	Succeeded.	Success	2025-03-27T09:20:19.88

Role Creation

category	correlationId	operationName	properties	result	resultType	time
Administrative	5cd3ea5f-c914-47b1	.../ROLEASSIGNMENTS/WRITE	{"requestbody":"{\"Id\":\"2c104baf-53b6-4s8e-8b82-f	Started.	Start	2025-03-31T17:03:19.28
Administrative	5cd3ea5f-c914-47b1	.../ROLEASSIGNMENTS/WRITE	{"statusCode":"Created","serviceRequestId":null,"eve	Succeeded.Created	Success	2025-03-31T17:03:23.19

- Two types of events – VM Creation & Role Creation
- # of Events – 2 vs 3
- Relevant data for each – in different part of the event!

Azure Multiplicity Example - VM

VM Creation

category	correlationId	operationName	properties	result	resultType	time
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	{"eventCategory":"Administrative","entity":"/subscriptions/b40c17b4-1193-2928-bc4a-			
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	a7fddee1f992/resourcegroups/MTG-AZREP100/providers/Microsoft.Compute/virtualMachines/CLDF-			
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	cag7de4d21-1749832072470","message":"Microsoft.Compute/virtualMachines/write","hierarchy":"02676fb2-6d			
			3-42d7-166b-b1ac31d89fa8/b40c17b4-1193-2928-bc4a-a7fddee1f992"}			

Role Creation

category	correlationId	operationName	properties	result	resultType	time
Administrative	5cd3ea5f-c914-47b1	.../ROLEASSIGNMENTS/WRITE	{"requestbody":"{\"Id\":\"2c104baf-53b6-4s8e-8b82-f	Started.	Start	2025-03-31T17:03:19.28
Administrative	5cd3ea5f-c914-47b1	.../ROLEASSIGNMENTS/WRITE	{"statusCode":"Created","serviceRequestId":null,"eve	Succeeded.Created	Success	2025-03-31T17:03:23.19

Minimal data in first event of VM creation: **`Started`**

Azure Multiplicity Example - VM

VM Creation

category	correlationId	operationName	properties	result	resultType	time
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	{"eventCategory":"Administrative","entity":"/subscript	Started.	Start	2025-03-27T09:18:31.22
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	{"statusCode":"Created","serviceRequestId":null,"responseBody":"{\"name\":\"CLDF-			
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	cag7de4d21-1749832072470\",\"id\":\"/subscriptions/b40c17b4-1193-2928-bc4a-			
			a7fddee1f992/resourceGroups/MTG-AZREP100/providers/Microsoft.Compute/virtualMachines/CLDF-			
			cag7de4d21-1749832072470\",\"type\":\"Microsoft.Compute/virtualMachines\",\"location\":\"westus\",\"tags			
			\":{\"Name\":\"Backup-TransientStormInstance\",\"bk_version\":\"9.1.3-p7-28390\",			
			\"bk_storm_instance_handle_id\":\"b755246b-3bd3-4298-a13a-7cd5770f4754\",\"bk_cluster_id\":			
			\"cb8ad3ea-0297-4c9d-9643-d0626beb54d9\",\"location_unique_id\":\"4b69f8be-13a1-4d6b-94ce-6432f4d4613			
			\",\"bk_instance_class\":\"TransientStormInstance\",\"bk_job_id\":			
			\"STORM_INSTANCE_RUN_aacd3688-058f-4ba4-b213-b8d36fda4c68:::0\",\"storm_type\":\"CLDF\",\"Departmer			
			\":\"Technology Services\",\"Division\":\"Security\",\"Owner\":\"John Doe\",\"Product\":\"Backup\"},\"properties			
			\":{\"hardwareProfile\":{\"vmSize\":\"Standard_DS4_v2\"},\"provisioningState\":\"Creating\",\"vmId\":\"087d5e15-			
			dc2a-4670-843c-28e8ebdd7c9a\",\"storageProfile\":{\"imageReference\":{\"id\":			
			\"/subscriptions/b40c17b4-1193-2928-bc4a-a7fddee1f992/resourceGroups/MTG-			
			AZREP100/providers/Microsoft.Compute/images/CLDF-baf7ee3d6b-osDisk-image-1743067078770\"},\"osDisk			
			\":{\"osType\":\"Linux\",\"name\":\"CLDF-baf7ee3d6b-osDisk-1743067078770\",\"createOption\":\"FromImage\"			
			\"caching\":\"ReadWrite\",\"managedDisk\":{\"storageAccountType\":\"Premium_LRS\"},\"diskSizeGB\":400},			
			\"dataDisks\":[{\"lun\":0,\"name\":\"CLDF-baf7ee3d6b-dataDisk-0-1743067078770\",\"createOption\":			
			\"FromImage\",\"caching\":\"ReadWrite\",\"managedDisk\":{\"storageAccountType\":\"Premium_LRS\"},			
			\"diskSizeGB\":256}]},\"osProfile\":{\"computerName\":\"CLDF-baf7ee3d6b\",\"adminUsername\":\"ubuntu\",			
			\"linuxConfiguration\":{\"disablePasswordAuthentication\":false},\"secrets\":[]},\"networkProfile			
			\":{\"networkInterfaces\":[{\"id\":\"/subscriptions/b40c17b4-1193-2928-bc4a-			

Role Creation

category	correlationId	operationName
Administrative	5cd3ea5f-c914-47b1	.../ROLEASSIGNMENTS/WRITE
Administrative	5cd3ea5f-c914-47b1	.../ROLEASSIGNMENTS/WRITE

Most data in 2nd event:
`Accepted.Created`

MITIGA

Azure Multiplicity Example - VM

VM Creation

category	correlationId	operationName	properties	result	resultType	time
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	{"eventCategory":"Administrative","entity":"/subscript	Started.	Start	2025-03-27T09:18:31.22
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	{"statusCode":"Created","serviceRequestId":null,"res	Accepted.Created	Accept	2025-03-27T09:18:33.31
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	{"eventCategory":"Administrative","entity":"/subscriptions/b40c17b4-1193-2928-bc4a-a7fddee1f992/resourcegroups/MTG-AZREP100/providers/Microsoft.Compute/virtualMachines/CLDF-cag7de4d21-1749832072470","message":"Microsoft.Compute/virtualMachines/write","hierarchy":"02676fb2-6d3-42d7-166b-b1ac31d89fa8/b40c17b4-1193-2928-bc4a-a7fddee1f992"}			

Role Creation

category	correlationId	operationName				
Administrative	5cd3ea5f-c914-47b1	.../ROLEASSIGNMENTS/WRITE	{"requestbody":"{\"Id\":\"2c104baf-53b6-4s8e-8b82-f	Started.	Start	2025-03-31T17:03:19.28
Administrative	5cd3ea5f-c914-47b1	.../ROLEASSIGNMENTS/WRITE	{"statusCode":"Created","serviceRequestId":null,"eve	Succeeded.Created	Success	2025-03-31T17:03:23.19

Again – Minimal data in last event: **Succeeded**

Azure Multiplicity Example - Role

VM Creation

category	correlationId	operationName
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE

Role Creation

category	correlationId	operationName
Administrative	5cd3ea5f-c914-47b1	.../ROLEASSIGNMENTS/WRITE
Administrative	5cd3ea5f-c914-47b1	.../ROLEASSIGNMENTS/WRITE

{"requestbody":"{\"Id\":\"2c104baf-53b6-4s8e-8b82-f0ac04d8db00\",\"Properties\":{\"PrincipalId\":\"2d33003f-cb1c-4ee6-a3f6-7188a3c51568\",\"PrincipalType\":\"Group\",\"RoleDefinitionId\":\"/providers/Microsoft.Authorization/roleDefinitions/acdd72a7-3385-48ef-bd42-f606fba81ae7\",\"Scope\":\"/subscriptions/b40c17b4-1193-2928-bc4a-a7fddee1f992/resourceGroups/DS-PROD-DataBricks-RG/providers/Microsoft.Databricks/workspaces/DS-PROD-Databricks\",\"Condition\":null,\"ConditionVersion\":null}}","eventCategory":"Administrative","entity":"/subscriptions/b40c17b4-1193-2928-bc4a-a7fddee1f992/resourceGroups/DS-PROD-DataBricks-RG/providers/Microsoft.Databricks/workspaces/DS-PROD-Databricks/providers/Microsoft.Authorization/roleAssignments/2c104baf-53b6-4s8e-8b82-f0ac04d8db00","message":"Microsoft.Authorization/roleAssignments/write","hierarchy":"02676fb2-6d33-42d7-16b-b1ac31d89fa8/b40c17b4-1193-2928-bc4a-a7fddee1f992"}

Most data in the first event: **`Started`**

Azure Multiplicity Example - Role

VM Creation

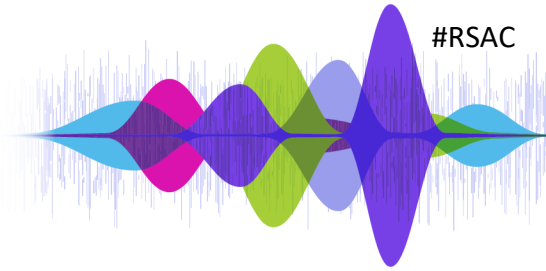
category	correlationId	operationName	properties	result	resultType	time
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	{"eventCategory":"Administrative","entity":"/subscript	Started.	Start	2025-03-27T09:18:31.22
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	{"statusCode":"Created","serviceRequestId":null,"res	Accepted.Created	Accept	2025-03-27T09:18:33.31
Administrative	b14b721d-9ce3-44f9	.../VIRTUALMACHINES/WRITE	{"statusCode":"Created","serviceRequestId":null,"eventCategory":"Administrative","entity":"/subscriptions/b40c1			

Role Creation

category	correlationId	operationName
Administrative	5cd3ea5f-c914-47b1	.../ROLEASSIGNMENTS/WRITE
Administrative	5cd3ea5f-c914-47b1	.../ROLEASSIGNMENTS/WRITE

7b4-1193-2928-bc4a-a7fddee1f992/resourceGroups/DS-PROD-DataBricks-RG/providers/Microsoft.Databricks/workspaces/DS-PROD-Databricks/providers/Microsoft.Authorization/roleAssignments/2c104baf-53b6-4s8e-8b82-f0ac04d8db00","message":"Microsoft.Authorization/roleAssignments/write","hierarchy":"02676fb2-6d33-42d7-16b-b1ac31d89fa8/b40c17b4-1193-2928-bc4a-a7fddee1f992"}

Minimal data in the second/last event: **`Succeeded.Created`**




```python
def azure_activity_successful_creation(azure_activity_df: DataFrame, successful_creation_status_column: str = ""):
    if successful_creation_status_column == "":
        successful_creation_status_column = "mtg_successful_creation"

    temp_df = azure_activity_df.groupBy("correlationId", "operationName", "resourceId").agg(
        sf.collect_set(sf.lower(sf.col("resultType"))).alias("result_types_set"),
        sf.collect_set(sf.lower(sf.col("resultSignature"))).alias("result_signatures_set"),
        sf.collect_set(sf.col('properties')).alias('properties')
    )

    # Pad the lists to be able to create a dict the correlates between result code and property value
    temp_df = temp_df.withColumn("mtg_max_length", sf.expr("GREATEST(size(result_signatures_set), size(properties))"))
    temp_df = temp_df.withColumn("result_signatures_set", pad_array_with_index("result_signatures_set", "mtg_max_length", "unknown"))
    temp_df = temp_df.withColumn("properties", pad_array_with_index("properties", "mtg_max_length", "None"))

    # Create a new column with a dictionary where keys are from 'result_signatures_set' and values are from 'properties'
    temp_df = temp_df.withColumn("properties_dict", sf.map_from_arrays(sf.col("result_signatures_set"), sf.col("properties")))\
                     .drop('properties', 'mtg_max_length')

    # join dataframes
    azure_activity_df = azure_activity_df.join(temp_df,
                                               on=["correlationId", "operationName", "resourceId"],
                                               how="left")

    azure_activity_df = azure_activity_df.withColumn(
        successful_creation_status_column,
        sf.when(
            (
                (~sf.array_contains(sf.col("result_types_set"), "failure"))
                & (
                    sf.array_contains(sf.col("result_signatures_set"), "accepted.created")
                    | sf.array_contains(sf.col("result_signatures_set"), "succeeded.created")
                )
            ),
            True,
        ).otherwise(False),
    )
    azure_activity_df = azure_activity_df.drop("result_types_set", "result_signatures_set")

    return azure_activity_df
```

```python
# Function to return if we had a successful
# creation operation in an azure activity
# correlation

# :param azure_activity_df: azure activity
# logs, filtered to a specific operation

# :param successful_creation_status_column:
# column name for status

# :return: the input DF with an extra column
# of successful_creation_status_column with
# True or False
```

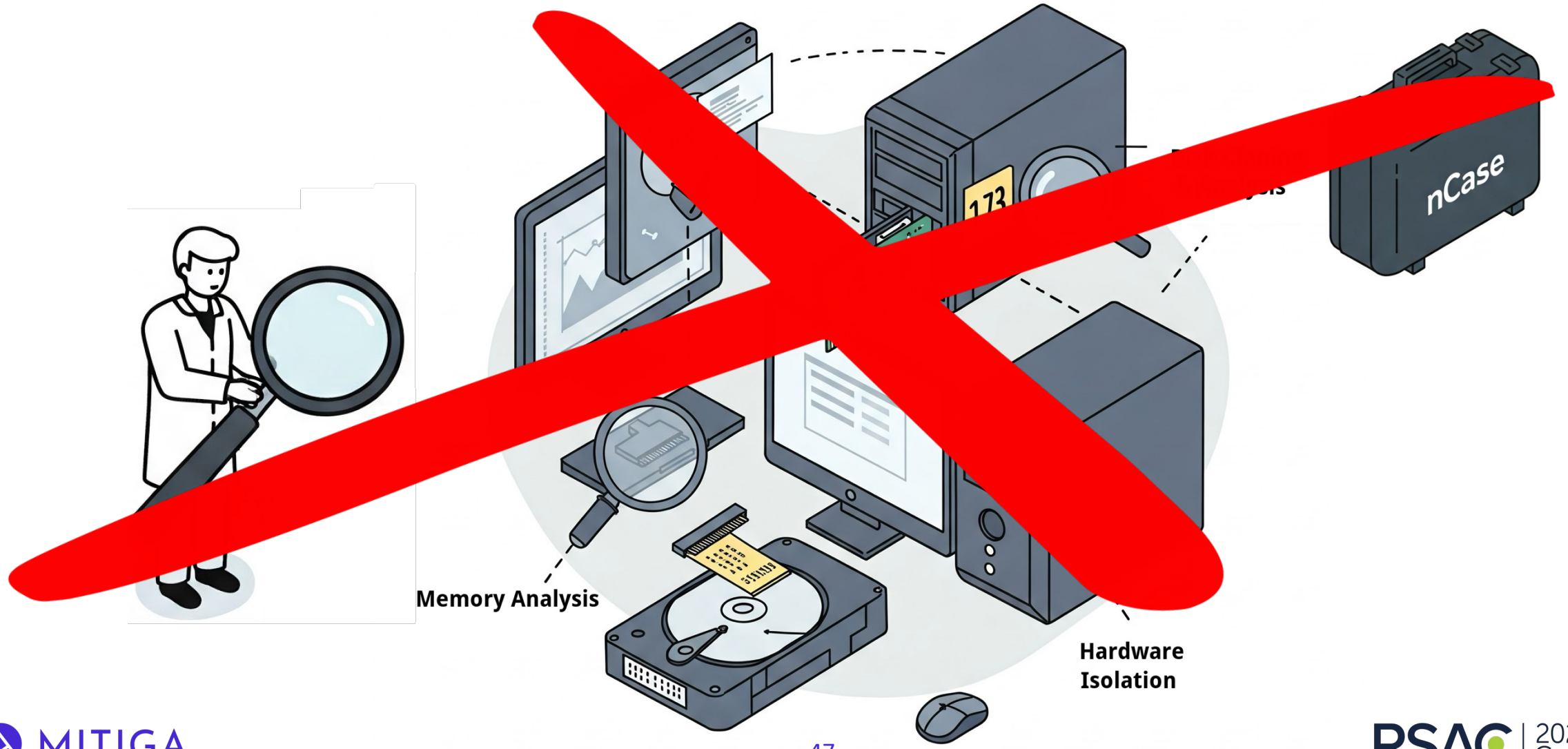
RSAC 2025 Conference

Forensically Sound

What Does That Mean in the Cloud?

Many Voices.
One Community.

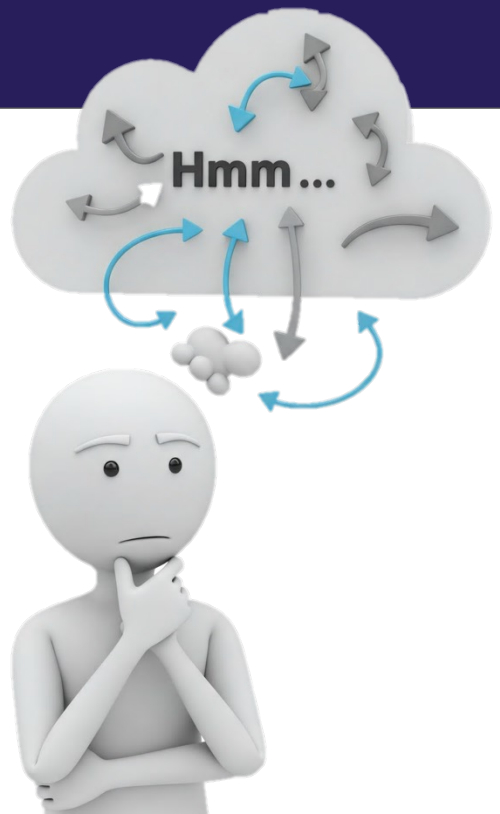
Forensically Sound Collection?



What Is Forensics Investigation?

Forensically sound evidence refers to digital evidence that has been collected, analyzed, handled, and stored in a manner that is **acceptable by the law**, and for which there is **reasonable evidence to prove** that the **integrity of the evidence** has been maintained.

Words of Wisdom by Gemini



- **Preservation of the Original**
- **Hashing/Signing**
- **Chain of Custody**
- **Write/Delete Prevention**
- **Proper Storage**

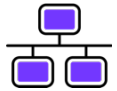
Originals, Hashing/Signing, Custody

What Is the Original?

- No way to preserve the Cloud/SaaS state
- Are logs originals? Or just artifacts?
- Are logs immutable?

What Should We Sign?

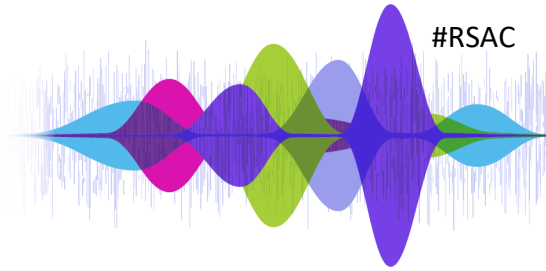
- Raw logs?
- Data lake/ warehouse logs?



Proving Chain of Custody

- All conducted via 3rd party cloud services and code
- Multiple signing/ hashing checkpoints?

Immutable Storage



1

Provided by All Cloud Providers

AWS Object Lock

GCP Bucket Lock/ Object Retention Lock

Azure Immutability Policies

2

Default for Security Log Archives

90 days cloud trail,
AWS Security Lake, etc.

3

Manually Defined for Everything Else

4

Can Be Used for Data Warehouse Storage As Well

RSAC 2025 Conference

Summary & Recommendations

Getting Ready for Your Next Forensics Investigation

Many Voices.
One Community.

Tomorrow, 2:25PM!

RSAC™ | 2025 Conference

Many Voices.
One Community.

SESSION ID: ANI-W09

It’s Getting Real & Hitting the Fan 2025 Edition

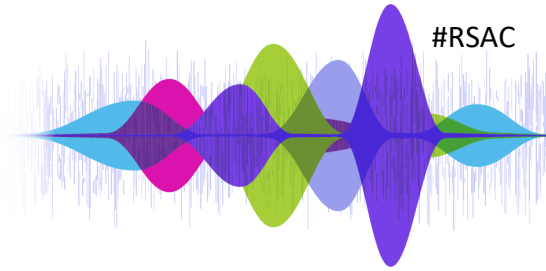
Think You See Me? No You Don’t!

Summary & Recommendations

1. Cloud forensics is a whole new world
2. **Cloud Telemetry Is Challenging!**
   Volume, structure, content, handling, retention and more…
3. Traditional approaches don’t scale well in cloud
4. **New Approach Needed!**
   Leveraging data warehousing and PySpark for advanced investigation

Don’t wait for your next breach to build your cloud forensics capabilities!

Apply What You Learned Today!



1

Next Week:

Map your gap in cloud forensics logs. Which logs do you have? How are you collecting them?
Do you have an infrastructure for cloud forensics investigation?

2

Next Month:

Start generating logs relevant for investigation. Set up the ability to collect and convert logs to a data warehouse and investigate it. Start familiarizing yourself with your logs structure.

3

And Then:

Build a continuous data warehouse collection approach. Create your automations. Practice (and find bad actors) with threat hunting with PySpark.
Repeat.

RSAC 2025 Conference

Many Voices.
One Community.

Thank You

Ofer Maor
CTO, Mitiga

linkedin.com/in/ofermaor/

Follow

MITIGA